Niveau(x) d'enseignement: 7 Numéro EDUQ : 5470124
Centre d'intérêt: 570

Cormier, Roger-M., 1932-.

La relation entre l'évaluation de l'enseignement par les étudiants et leur rendement académique individuel ou collectif : un dépouillement des écrits / par Roger M. Cormier, collab. de Richard Brunet, Jean-Marc Dumas, Françoys Gagné et Ghislaine Lizotte. -- [Montréal] : INRS-Éducation, [197A?].
14, [4] f. ; 28 cm.

Document: 0574-01. Bibliogr. : f. [15-18].

Descripteur(s): *Evaluation de l'enseignant *Evaluation par l'étudiant *Satisfaction de l'étudiant *Attitude de l'enseignant *Relations maître-élèves *Evaluation de l'enseignement *Rendement scolaire *Test *Revue de la documentation *Collégial (Niveau)

Identificateur(s) de nom: *PERPE (Test)

La relation entre le rendement scolaire des étudiants et leur évaluation de l'enseignement a fait l'objet de nombreuses études dont les résultats apparaissent à première vue contradictoires. Le document consiste en une revue de la documentation publiée sur le sujet. On y distingue les études selon qu'elles prennent les étudiants ou les professeurs comme unités d'analyse. Dans la première partie, l'auteur analyse les études qui ont mis en rapport l'évaluation de l'enseignement par les étudiants avec le rendement scolaire individuel de ceux-ci. Dans la deuxième partie, il analyse les études qui ont mis en rapport l'évaluation de l'enseignement par les étudiants avec la libéralité du professeur dans l'attribution des notes et avec l'évaluation objective du rendement scolaire de chaque groupe d'étudiants.

L'étude s'appuie sur une bibliographie d'une cinquantaine de titres, pour la plupart de langue anglaise, généralement postérieurs à 1956.

Type de document: r

EDUQ 8470124

INRS-Education

La relation entre l'évaluation de l'enseignement par les étudiants et leur rendement académique individuel ou collectif

Un dépouillement des écrits

par

Roger A. Cormier*

Document 0574-01

* L'auteur désire souligner la collaboration de Richard Brunet, Jean-Marc Dumas, François Gagné et Ghislaine Lizotte dans la réalisation de ce travail. Il les en remercie sincèrement.

De nombreuses études ont été consacrées au problème de la relation entre le rendement académique et l'évaluation de l'enseignement par les étudiants. Déjà, Remmers en 1928, constatait que le rendement académique de l'étudiant à l'intérieur d'une classe pouvait aussi bien influencer défavorablement que favorablement l'évaluation qu'il fait de l'enseignement. Par la suite, des études vinrent tantôt confirmer ces résultats, tantôt les contredire.

Dans une analyse des écrits concernant la validité des questionnaires d'évaluation des professeurs par les étudiants, Costin, Greenough et Menges (1971) concluent qu'un grand nombre de chercheurs n'ont trouvé aucune relation entre les notes scolaires et l'évaluation des professeurs par les étudiants, et que lorsqu'on trouvait une telle relation, elle était plutôt faible. Pourtant, Anikeef (1953) que citent Costin et al. a trouvé dans trois groupes de professeurs, des corrélations respectives de 0.73 (N = 13), 0.43 (N = 17) et 0.51 (N = 19) entre la distribution des scores d'évaluation de l'enseignement par les étudiants et celle des moyennes des notes attribuées par les professeurs dans chacune de leur classe. Plus récemment, Rodin et Rodin, observaient chez des professeurs adjoints une relation négative très élevée (r =-.75) entre l'évaluation de l'efficacité de l'enseignement par les étudiants et leur progrès académique.

Notons que dans ces deux dernières études, l'unité d'analyse est la classe ou le professeur alors que dans plusieurs études, l'unité d'analyse est l'étudiant. Morsh et Wilder (1954) ont souligné l'importance de faire cette distinction. A notre connaissance, seul Chabot, et ce brièvement, a étudié les deux phénomènes simultanément. Avant de reprendre l'étude de Chabot plus à fond avec les données des nouvelles formes expérimentales du questionnaire PERPE (automne 1973), nous avons procédé à une analyse des écrits sur le sujet. C'est l'objet de ce document. L'étude empirique elle-même fera l'objet d'une communication qui sera présentée au

congrès de la SCEE le mois prochain à Toronto (Cormier et Gagné, 1974).

Dans l'analyse qui suit, nous distinguons les études selon qu'elles prennent les étudiants ou les professeurs comme unités d'analyse. Dans la première partie, nous analyserons donc les études qui ont mis en rapport l'évaluation de l'enseignement par les étudiants avec le rendement académique individuel de ceux-ci. Dans la deuxième partie, nous analyserons les études qui ont mis en rapport l'évaluation de l'enseignement par les étudiants avec, d'une part, la libéralité du professeur dans l'attribution des notes (grading linency), et, d'autre part, l'évaluation objective du rendement académique de chaque groupe d'étudiants.

Le rendement académique des étudiants pris individuellement

De nombreux chercheurs ont étudié la relation qui existe entre le rendement académique des étudiants pris individuellement et l'évaluation qu'ils font de l'enseignement. Il faut cependant distinguer dans ces études celles qui mettent ces deux variables en relation à l'intérieur de plusieurs classes (e.g. à l'intérieur de 20 classes de 30 étudiants) de celles qui regroupent les données de plusieurs classes (e.g. les données de 600 étudiants provenant de 20 classes).

Etude à l'intérieur de mêmes classes

Remmers (1928 et 1930) fut l'un des premiers à s'intéresser à la relation que l'on peut observer entre le rendement académique des étudiants et l'évaluation qu'il fait de l'enseignement. Il administra le Purdue Rating Scale for Instructors dans 17 classes et demanda à un certain nombre d'étudiants, i.e. ceux dont les résultats scolaires étaient supérieurs, de faire un X sur leur feuille d'évaluation. Il obtint des coefficients de corrélation bi-sérial variant de -.86 à .89 avec une moyenne de .05. Ces coefficients furent calculés sur chacun des 10 items du questionnaire ainsi que sur le score total. La moyenne des coefficients des 17 classes ($15 < N$ étudiants > 33)

sur le score total était .07 (les r variant de -.14 à .31). Cette moyenne ne diffère donc pas énormément de l'autre. On ne peut donc supposer qu'un faible degré de fidélité* au niveau de chaque item aurait pu atténuer les coefficients de corrélation. Smeltzer et Harter (1934) observèrent des résultats analogues dans 12 classes à l'aide du score total d'un questionnaire de 42 items. Les coefficients varièrent de -.20 à .18 ($\bar{r}$ = .02) sur une évaluation anonyme et de -.14 à .17 ($\bar{r}$ = .03) sur une évaluation signée par les étudiants. Bowman (1934) de son côté obtint dans neuf classes sur une évaluation globale à l'aide d'un seul item, des coefficients variant de .00 à .65 ($\bar{r}$ = .26).

Chabot (1970), lors de l'expérimentation du questionnaire PERPE, demanda aux étudiants de 88 classes d'indiquer leurs résultats scolaires à l'intérieur de l'une ou l'autre de 6 catégories (Moins de 60% à 80% et plus), résultats qu'elle mit en relation avec le score global d'insatisfaction de chaque étudiant. Elle obtint dans 10 classes des coefficients de corrélation** variant de -.20 à +.09 avec une moyenne de -0.05. Chabot effectua également une analyse du rendement collectif des étudiants, analyse que nous étudions plus bas.

Les autres études (Bendig, 1953r; Garverich et Carter, 1962; Guthrie, 1954; Holmes, 1972; Quereshi & Widlak, 1971; Schuk & Crivelli, 1973; Treffinger et Feldhusen, 1970) que nous avons retracé concernant la relation entre le rendement académique des étudiants à l'intérieur d'une classe et leur évaluation de l'enseignement n'utilisent que 5 classes au moins.

---

* Dans de nombreux cas, les chercheurs ne rapportent pas de coefficients de précision des instruments qu'ils utilisent. Etant donné que le degré de précision de deux variables influence directement leur coefficient de corrélation (voir Cormier et Gagné, 1973), nous avons tenu compte de ce facteur dans notre analyse. Nous n'en ferons plus mention à moins de douter fortement du degré de précision des instruments utilisés.

** En ajustant ses résultats selon un continuum insatisfaction-satisfaction. A moins d'indication contraire, les résultats que nous rapportons dans notre analyse se réfère à un continuum faible-fort sur les deux variables.

Malgré les résultats de Bowman, et compte tenu des résultats obtenus par Richard Brunet*, nous nous croyons justifiés de conclure que le rendement académique de l'étudiant à l'intérieur de la classe influencerait l'évaluation qu'il fait de l'enseignement tantôt d'une façon favorable, tantôt d'une façon défavorable, l'influence favorable se présentant possiblement un peu plus souvent, surtout s'il s'agit des étudiantes (voir Quereshi & Widlak, 1971; Bendig, 1953) ainsi que Treffinger et Feldhusen, 1970).

Regroupement de données provenant de classes différentes

Un grand nombre d'études concernant la relation entre le rendement académique des étudiants pris individuellement et l'évaluation qu'ils font de l'enseignement porte sur des données provenant de plusieurs classes groupées ensemble. L'interprétation de ces données est très difficile à faire, particulièrement lorsqu'on n'utilise pas un coefficient de corrélation, étant donné que le regroupement des étudiants augmente la puissance des tests d'inférence statistique. Par exemple, Stewart et Malpass (1966), en regroupant les données de 1975 sujets, rapportent des chi-carré significatifs à un niveau de probabilité d'erreur de 1/10 de 1% et concluent à une relation très étroite entre les deux variables. Si nous convertissons ces chi-carrés en coefficients de contingence, nous obtenons des coefficients variant de 0.17 à 0.28**. Painter et Granzin (1972) arrivent à des résultats analogues.

D'autres chercheurs divisent les étudiants en groupes selon leurs notes (e.g. A, B, C, D et E) et utilisent un test t ou l'analyse de variance pour déterminer s'il y a une différence entre les groupes. Ainsi, Weaver (1960) rejette au moyen de tests t l'hypothèse nulle à un niveau de probabilité d'erreur allant jusqu'à 1/10 de 1% entre certains groupes sur des données que nous présentons au tableau 1. D'après les données de ce tableau, rien ne

---

* Communication verbale concernant un $\overline{r} = .00$

** Les détails que nous donnent les auteurs nous ont également permis de calculer un r de Pearson. Pour un C de 0.28 nous obtenons un r de 0.30.

Tableau 1

Score moyen de l'évaluation de l'enseignement selon la note que l'étudiant s'attend de recevoir (d'après les données de Weaver, 1960, p.23)

	A N=31	B N=268	C N=370	D N=30
Score total*	96.10	94.56	91.15	84.63
σ	(13.91)	(12.[illegible])	(11.51)	(21.33)
Personnalité	26.45	27.27	26.60	27.40
σ	(4.27)	(3.57)	(2.60)	(11.91)
Méthode d'enseignement*	69.94	67.36	65.14	61.47
σ	(11.07)	(9.82)	(6.36)	(11.15)

* $P < .001$ entre B + C et (A + B) - (C + D) sur des tests t.

nous permet de croire que l'on obtiendrait des coefficients de corrélation très élevés, si nous utilisions cette technique. Richard Brunet à l'aide de la technique de la simulation sur ordinateur a estimé que le coefficient de corrélation entre le rendement académique et le score total serait de 0.12. Si nous calculons des r (voir Cohen,1969) sur chacun des t que l'auteur a trouvé sur le score global nous obtenons des résultats variant de .03 à .31.

Bausell et Magoon (1972), Bendig (1952), Bledsoe, Brown et Strickland (1971), Echandia (1964), Kennedy (1972), Russell (1951), Russell et Bendig (1953) ainsi que Walker (1968), à l'aide de l'analyse de variance ou du test t, arrivent à des résultats semblables. Notons que Bledsoe, Brown et Strickland trouvèrent une interaction statistiquement significative (N = 4368) entre le sexe de l'étudiant et les notes de cours sur 2 des 6 facteurs de l'évaluation de l'enseignement. Seul Blum (1936) conclut à l'absence de relation entre les deux variables et ce à l'aide de 57 étudiants provenant de 2 classes seulement.

Dans un certain nombre d'études, les auteurs ont utilisé la technique du coefficient de corrélation, ce qui nous donne une meilleure idée

de l'importance du phénomène. Nous présentons ces coefficients au tableau 2. Notons que dans trois des quatre études rapportées, nous croyons que l'indice de précision du score de l'évaluation de l'enseignement est probablement faible, étant donné qu'on a calculé des coefficients de corrélation au niveau des items alors que l'unité d'analyse est l'étudiant. Des coefficients corrigés pour l'atténuation serait probablement de 1.5 à 2 fois supérieurs à ceux indiqués au tableau 2.

Nous avons également retracé deux autres études (Tolor, 1973 et Riley et Lifshitz, 1950) dans lesquelles on regroupe les données des étudiants de plusieurs classes mais que nous ne pouvons placer dans l'un ou l'autre des groupes déjà cité dans cette section à cause des particularités des techniques qu'ils utilisent. Il semble bien que les résultats de ces études ne s'éloignent pas de ceux que nous avons déjà présentés. Notons que Riley & Lifshitz obtiennent une relation beaucoup plus prononcée pour l'item relié à l'impartialité concernant les examens (p. 84).

Tableau 2

Coefficient de corrélation entre différentes distributions de scores de rendement académique (réel ou attendu) et de scores d'évaluation de l'enseignement

Auteurs	Nombre		Coefficient
	Etudiants	Classes	
Caffrey (1969)*	131	3	$.23 < r > .32$ sur 6 items $r > .23$ sur 36 items
Hudelson (1951)*	192		$r = 0.19$
Rayder (1968)*	4032	97	$-0.03 < r > 0.18$
Starrak**	des milliers		$\bar{r} = .15$

* Degré de précision de l'évaluation de l'enseignement probablement faible.

** Aucune idée quant à l'indice de précision.

Les études que nous avons analysées dans cette section nous permettent de conclure que lorsque l'on groupe les données des étudiants de plusieurs classes, il existe un lien entre les notes des étudiants et l'évaluation qu'ils font de l'enseignement. Si nous tentons de chiffrer ce lien en termes de coefficient de corrélation de Pearson et compte tenu de la correction pour l'atténuation, nous constatons qu'il se situe généralement entre -.05 et 0.50 et nous pouvons dire sans trop grand risque d'erreur que la moyenne doit se situer près de 0.25.

Cette constatation contredit-elle celle que nous avons faite dans la section précédente? Pas du tout. Le regroupement des données provenant de plusieurs classes introduit un autre élément qui peut influencer les résultats. Il s'agit de la sévérité du professeur dans l'attribution de ses notes. Nous étudions ce facteur dans la section qui suit.

Le rendement académique des étudiants pris collectivement

Dans les deux sections précédentes, nous avons rapporté des articles dont l'unité d'analyse était l'étudiant. D'autres chercheurs, dans le domaine de l'évaluation de l'enseignement, ont pris la classe comme point de départ de leur étude.

Si l'on met en relation l'évaluation de l'enseignement par les étudiants avec la moyenne des notes de classes des mêmes étudiants, deux phénomènes passablement distincts peuvent se présenter. Si les professeurs préparent eux-mêmes les examens et s'ils sont libres de fixer la moyenne pour leur classe, nous analyserons en toute probabilité le lien entre l'évaluation de l'enseignement par les étudiants et la libéralité du professeur dans l'attribution des notes (grading leniency). Par contre si le professeur n'a aucun contrôle sur l'évaluation du rendement académique des étudiants, nous sommes alors en présence d'une étude de la relation entre l'évaluation de l'enseignement par les étudiants et une évaluation opérationnalisée du rendement du professeur. Ce sont ces phénomènes que nous analyserons dans les deux sections qui suivent.

La sévérité du professeur

Heilman et Armentrout (1936) n'ont trouvé dans 50 classes aucun lien (r = -.04) entre une évaluation globale de l'enseignement par les étudiants et la générosité du professeur. Ils obtiennent cependant une corrélation négative de 0.24 sur l'item "Fairness in Grading". Ainsi les classes d'étudiants dont la note est inférieure ont tendance à évaluer leur professeur plus favorablement à l'égard de leur impartialité dans l'attribution des notes.

Jiobu et Pollis (1971) concluent également à une absence de lien entre l'évaluation de l'enseignement par les étudiants dans 67 classes et la générosité des professeurs. Ils obtiennent sur deux items des coefficients de corrélation de 0.18 (la perception de l'apprentissage) et 0.30 (évaluation globale) mais ces coefficients ne sont pas statistiquement significatifs à un niveau de probabilité d'erreur de 1%. Guthrie (1954), sans y faire allusion, semble avoir étudié le même phénomène. Il calcula un coefficient de corrélation entre la distribution des moyennes des notes de 77 classes et celle des scores d'une évaluation par les étudiants. Il obtint un r de .07.

Contrairement aux trois études précédentes, Anikeef (1953) et Chabot (1970) ont trouvé un lien positif entre les deux mêmes dimensions. Le premier a obtenu des coefficients de corrélation de 0.73 (N = 13), 0.43 (N = 17) et 0.51 (N = 19) selon les niveaux d'enseignement collégial. La deuxième a obtenu un coefficient de corrélation de 0.42 dans 88 classes de CEGEP. En utilisant la moyenne de la classe pour l'ensemble des matières, ce coefficient tombe à 0.14.

Voeks et French (1960) concluent à l'absence de relation entre l'évaluation de l'enseignement par les étudiants et la générosité des professeurs. Cependant dans un des trois départements où ils recueillent des données, ils obtiennent un coefficient de corrélation de 0.60 (N = 11 et $p > .05$) entre les deux variables. Dans une autre analyse, ils constatent que dans 4 départements le professeur le mieux coté par les étudiants attribue des notes statistiquement plus élevées ($p < .01$) que le professeur le

moins bien coté, que dans cinq autres départements le même phénomène persiste quoique non statistiquement significatif mais que dans un dernier département (en économique) le contraire se produit et ce avec une différence statistiquement significative ($p < .01$).

Ces six études nous permettent de conclure à l'existence d'un lien plus ou moins important entre l'évaluation de l'enseignement et la générosité des professeurs dans l'attribution des notes. Ce lien varierait selon les dimensions mesurées (Heilman & Armentrout) et selon les matières enseignées (Voeks et French).

Cette relation entre l'évaluation de l'enseignement par les étudiants et la générosité des professeurs a fort probablement une influence sur la relation qui a été observée dans la section précédente. En regroupant les données des étudiants provenant de plusieurs classes, les étudiants dont la moyenne de classe est plus élevée auront tendance à évaluer leur professeur plus favorablement que ceux dont la moyenne de classe est moins élevée. Avant de faire de tels regroupements il y aurait donc lieu de transformer le rendement scolaire de chaque étudiant en score z à l'intérieur de chaque classe. C'est ce que nous ferons dans notre étude empirique.

Le rendement des professeurs

Nos données ne nous permettrons pas de vérifier le lien qu'il y aurait entre l'évaluation de l'enseignement par les étudiants et le rendement des professeurs tel que mesuré par le rendement académique des étudiants. Une analyse des travaux dans lesquels on aborde ce problème nous apparaît nécessaire afin que nous puissions en tenir compte dans les conclusions de notre recherche empirique.

La plupart des chercheurs, sauf Bendig (1953a) et Rodin & Rodin (1972) ont trouvé un lien positif entre l'évaluation de l'enseignement par les étudiants et le rendement des professeurs mesuré opérationnellement de diverses façons.

Les résultats de Bendig ne proviennent d'un groupe de cinq professeurs et semblent contraires à ceux qu'on a trouvé dans la même institution (Bendig, 1953r, Russell, 1951 et Russell et Bendig, 1953).

L'étude de Rodin et Rodin porte sur un plus grand nombre de cas. Il s'agit qu'un cours magistral donné trois fois par semaine par un professeur avec en plus 2 séances de laboratoires (small recitation sections) sous la responsabilité de 12 assistants. Le rendement des étudiants ayant été ajusté en fonction de leur niveau de connaissances au début du cours, Rodin et Rodin obtinrent un coefficient de corrélation de -0.75 entre le rendement moyen par groupe d'étudiants et l'évaluation de l'enseignement par les mêmes groupes d'étudiants. Nous avons retracé trois études (Frey, 1973, Gessner, 1973 et Sullivan et Skanes, 1973) qui infirment ou expliquent les résultats de Rodin et Rodin.

Gessner dans une étude auprès d'une classe de 119 étudiants en médecine a constaté que les professeurs responsables des sujets où les étudiants ont le mieux réussi à un examen standardisé au niveau national étaient plus favorablement évalués par les étudiants. Il obtient des coefficients de corrélation de 0.77 (contenu et organisation du cours) et 0.69 (présentation). Gessner explique la contradiction entre ses résultats et ceux de Rodin et Rodin par la présence d'assistants professeurs dans cette dernière étude et la confusion qui peut régner entre leur évaluation et celle du responsable du cours.

De leur côté, Sullivan et Skanes obtiennent un coefficient de corrélation moyen de 0.39 dans dix groupes de classes dont les examens étaient préparés collectivement à l'intérieur de chaque groupe. De plus, à l'intérieur du groupe le plus nombreux (40 classes en psychologie) ils obtiennent, en séparant les professeurs à plein temps des professeurs adjointe au niveau du 2è et du 3è cycle, des coefficients de corrélation de 0.53 et .01 respectivement. Cette différence dans les coefficients de corrélation entre professeurs à plein temps et professeurs adjoints nous suggèrent que les résultats de Rodin et Rodin ne sauraient être appliqués aux professeurs à plein temps.

Enfin Frey calcula des coefficients de corrélation chez deux groupes de professeurs en mathématiques entre la distribution des scores d'un test objectif (corrigé pour tenir compte des aptitudes individuelles des étudiants) et celle de divers scores factoriels de l'évaluation des étudiants. Il obtint des coefficients variant de 0.14 à 0.91 chez les 8 professeurs du cours d'introduction et variant de 0.36 à 0.90 chez les 5 professeurs du cours plus avancé. Notons cependant que le questionnaire d'évaluation de l'enseignement fut administré après que les étudiants eurent pris connaissance de leurs notes. Le contrôle que Frey fait de ce facteur au moyen de la régression multiple et ce sans tenir compte de la sévérité des professeurs, nous apparaît discutable.

Dans chacune des trois études que nous venons de citer les autres font une critique assez poussée des résultats de Rodin et Rodin. Nous avons également retracé 7 autres études concernant le lien entre l'évaluation de l'enseignement par les étudiants et le rendement des professeurs. Nous constatons d'après le tableau 3 dans lequel nous résumons ces études que dans l'ensemble il y a un lien positif entre les deux variables. Ce lien n'est cependant pas aussi fort que ceux que l'on a constaté dans les trois études récentes que nous venons de résumer. Il y a même plusieurs coefficients de corrélation négatifs.

Soulignons que les études que nous présentons au tableau 3 ont été effectuées dans la plupart des cas chez des professeurs adjoints tout comme celle de Rodin et Rodin. Ces résultats ne sauraient donc s'appliquer à des professeurs de carrière comme nous l'avons indiqué plus haut.

Les études que nous venons de citer dans cette dernière section nous permettent donc d'avancer l'hypothèse que les étudiants peuvent être de bons évaluateurs de l'enseignement pour les professeurs à plein temps lorsque le rendement des étudiants est évalué par une mesure qui n'est pas préparée par les professeurs eux-mêmes.

Tableau 3

Résultats obtenus dans diverses études concernant le lien entre l'évaluation de l'enseignement et le rendement des professeurs

Etudes	Enseignants	Niveau	Matière enseignée	Critère d'évaluation de l'enseignement	Critère d'évaluation du rendement du prof.	Coefficients de corrélation
Elliott (1950)	36 adjoints sous la direction d'un seul professeur	Universitaire	Chimie	12 items pour les laboratoires et 12 items pour les "récitations"	Scores de tests objectifs ajustés pour l'aptitude	$\bar{r}$ = .24 (variant de -.08 à .43)
Lens (1946)	11 institutrices	Secondaire	Divers	Un seul item global	Gain scolaire	.06
McKeachie, Lin et Mann (1971)	Entre 6 et 17 professeurs dans 5 études	Universitaire	Psycho. Français Economique	4 ou 6 facteurs	Scores de tests objectifs ajustés pour l'aptitude	$\bar{r}$ = .09 (variant de -.60 à .72)
Morsh, Burgess et Smith (1956)	121 instructeurs	Cours de 8 jours dans l'aviation		Un seul item global 4 items spécifiques	Test écrit et test de performance	.40 r variant de -.40 à .46
Remmers, Martin, Elliott (1949)	37 adjoints sous la direction d'un seul professeur	Universitaire	Chimie	Purdue (24 items)	Scores de tests objectifs ajustés pour l'aptitude	Les meilleurs professeurs sont cotés plus favorablement sur 23 des 24 items
Rubinstein et Mitchell (1970)	24 adjoints	Universitaire	Psycho.	4 facteurs du Purdue	Scores de tests objectifs et de tests des assistants	r variant de -.22 à .44
Wherry (1952)	11 adjoints	Universitaire	Psycho.	Divers instruments	Scores de tests objectifs ajustés pour l'aptitude	$\bar{r}$ = .20 (variant de -.02 à .51)

Résumé

La relation entre le rendement académique des étudiants et leur évaluation de l'enseignement a fait l'objet de nombreuses études dont les résultats apparaissent à première vue contradictoires. Ces contradictions viennent, il nous semble, du fait qu'on n'a pas fait de distinction entre les différents types d'analyses qu'on peut effectuer dans ce domaine.

Une première distinction à faire est celle de l'unité d'analyse, selon que l'on met en relation des données individuelles ou collectives des étudiants. De plus, lorsque l'on utilise l'étudiant comme unité d'analyse, il faut distinguer les travaux où l'on étudie la relation entre le rendement académique et l'évaluation de l'enseignement à l'intérieur des différentes classes de ceux où l'on groupe ensemble les données des étudiants provenant de différentes classes. Enfin lorsqu'on étudie les moyennes de classe, il faut se demander si les professeurs ont un contrôle sur l'évaluation du rendement des étudiants.

Dans les études où l'on a étudié la relation entre le rendement académique et l'évaluation de l'enseignement à l'intérieur de différentes classes d'étudiants, l'on constate que cette relation peut être aussi bien négative que positive, la relation positive se présentant un peu plus souvent.

Dans les études où l'on a groupé les données d'étudiants provenant de plusieurs classes, la relation est toujours positive (sauf dans une étude utilisant 2 classes seulement) et dans l'ensemble la moyenne des coefficients de corrélation seront près de 0.25.

Ces derniers résultats semblent contredire ceux du premier groupe d'études, ce qui n'est pas le cas, il nous semble, si l'on tient compte d'un autre facteur, i.e. de la relation entre la sévérité des professeurs dans

l'attribution des notes scolaires et l'évaluation de l'enseignement par les étudiants, relation qui se manifeste d'après les résultats d'un autre groupe d'études où l'analyse s'est effectuée au niveau des classes d'étudiants ou des professeurs. Les étudiants ont tendance à évaluer plus sévèrement les professeurs qui sont plus sévères dans la correction des travaux et des examens. Ce facteur augmenterait artificiellement le coefficient de corrélation dans les études où l'on a groupé les données individuelles des étudiants provenant de plusieurs classes.

Enfin dans un dernier type d'analyse prenant également la classe ou le professeur comme unité d'analyse, l'on constate chez les professeurs à plein temps qu'il y a une relation positive assez importante entre l'évaluation de l'enseignement par les étudiants et le progrès de ces derniers tel que mesuré par des tests de rendement objectif ou standardisés. Cette relation cependant peut être aussi bien négative que positive dans le cas de professeurs à temps partiel.

Références

ANIKEEFF, Alexis M. (1953) Factors affecting student evaluation of college Faculty members. The Journal of Applied Psychology, 37, 458-460.

BAUSELL, R. Barker & MAGOON, Jon (1972) Expected grades in a course, grade point average, and student ratings of the course and the instructor. Educational and Psychological Measurement, 32, 1013-1023.

BENDIG, A.W. (1952) A preliminary study of the effect of academic level, sex, and course variables on student ratings of psychology instructors. Journal of Psychology, 34, 21-26.

BENDIG, A.W. (1953r) The relation of level of course achievement to students' instructor and course ratings in introductory psychology. Educational and Psychological Measurement, 5, 437-448.

BENDIG, A.W. (1953s) Student achievement in introductory psychology and student ratings of the competence and empathy of their instructors. The Journal of Psychology, 36, 427-433.

BLEDSOE, Joseph C., BROWN, Iva D., STRICKLAND, Arthur D. (1941) Factors related to pupil observation reports of teachers and attitudes toward their teacher. The Journal of Educational Research, 65, 119-126.

BLUM, M.L. (1936) An investigation of the relation existing between students' grades and their ratings of the instructors' ability to teach. Journal of Educational Psychology, 27, 217-221.

BOWMAN, E.C. (1934) Pupil ratings of student teachers. Educ. Adm. Superiors, 20, 141-146.

CAFFREY, Bernard (1969) Lack of bias in student evaluations of teachers. Proceedings, 77th Annual Convention, APA.

CHABOT, Marie (1970) La validité du test PERPE. Montréal: Service PERPE.

COHEN, Jacob (1969) Statistical power analysis for the behavioral sciences. New-York: Academic Press.

CORMIER, Roger A. & GAGNE, Françoys (1974) Some relationships between students grades, teachers leniency in grading and student satisfaction toward teaching. Communication devant être présentée au congrès annuel de la SCEE à Toronto.

CORMIER, Roger A. & GAGNE, Françoys (1973) La stabilité de certaines variables du questionnaire PERPE. Communication présentée au congrès annuel de l'ACFAS à Montréal.

COSTIN, Frank, GREENOUGH, William T. & MENGES, Robert J. (1971) Student ratings of college teaching: reliability, validity and usefulness. Review of Educational Research, 41, 511-535.

ECHANDIA, P.P. (1964) A methodological study and factor analytic validation of forced-choice performance ratings of college accounting instructors. Dissertation Abstracts, 25(4), 2605-2606.

ELLIOTT, D.H. (1950) Characteristics and relationships of various criteria of college and university teaching. Purdue University Studies in Higher Education, 70, 5-61.

FREY, P.W. (1973) Student instructional ratings and faculty performance. Communication présentée au congrès de l'AERA à New Orleans.

GARVERICK, Charles M. & CARTER, Harold D. (1962) Instructor ratings and expected grades. Californian Journal of Educational Research, 13, 218-221.

GESSNER, Peter K. (1973) Evaluation of instruction. Science, 180-566-570.

GUTHRIE, E.R. (1954) The evaluation of teaching: a progress report. Seatle: University of Washington.

HEILMAN, J.D. & ARMENTROUT, W.D. (1936) The rating of college teachers on ten traits by their students. The Journal of Educational Psychology, 27, 197-216.

HOLMES, David S. (1972) Effects of grades and disconfirmed grade expectancies on students' evaluations of their instructor. Journal of Educational Psychology, 63, 130-133.

HUDELSON, Earl (1951) The validity of student rating of instructors. School and Society, 73, 265-266.

JIOBU, Robert M. & POLLIS, Carol A. (1971) Student evaluations of courses and instructors. The American Sociologist, 6, 317-321.

KENNEDY, W. (1972) The relationships of selected student characteristics to components of teachers course evaluation among freshman english students at Kent State University. Paper read at the AERA Annual Meeting, Chicago.

LINS, L.J. (1946) The prediction of teaching efficiency. J. Exp. Educ., 15, 2-60.

MCKEACHIE, W.J., LIN, Y. & MANN, W. (1971) Students ratings of teacher effectiveness: validity studies. American Educational Research Journal, 8, 435, 445.

MORSH, Joseph E., BURGESS, George G. & SMITH, Paul N. (1956) Student achievement as a measure of instructor effectiveness. Journal of Educational Psychology, 47, 79-88.

MORSH, Joseph E. & WILDER, Eleanor W. (1954) Identifying the effective instructor: a review of the quantitative studies, 1900-1952. Document ERIC ED 044 371

PAINTER, John J., GRANZIN, Kent L. (1972) Consistency theory as an exploration of students' course evaluation tendencies. The Journal of Experimental Education, 41, 78-81.

QUERESHI, M.Y. & WIDLAK, Frederich W. (1971) Teacher evaluation as a function of the students' sex and achievement level. ERIC 059 972

RAYDER, Nicholas F. (1968) College student ratings of instructors. The Journal of Experimental Education, 37, 76-81.

REMMERS, H.H. (1928) The relationship between students' marks and student attitude toward instructors. School and Society, 2, 759-760.

REMMERS, H.H. (1930) To what extent do grades influence student ratings of instructors? Journal of Educational Research, 21, 314-316.

REMMERS, H.H., MARTIN, F.D. & ELLIOTT, D.N. (1949) Are students ratings of instructors related to their grades? Purdue University Studies in Higher Education, 66, 17-26.

RILEY, J., RYAN, B. & LIFSHITZ, M. (1950) The student looks at his teacher. New Brunswich, N.J.: Rutgers University Press.

RODIN, Miriam, RODIN, Burton (1972) Student evaluations of teachers. Student rate most highly instructors from whom they learn least. Science, 177, 1164-1166.

RUBINSTEIN, J., MITCHELL, H. (1970) Feeling free, student involvement and appreciation. Proceedings of the 78th Annual Convention of the American Psychological Association, 5, 623-624.

RUSSELL, H.E. (1951) Interrelations of some indices of instructor effectiveness: An exploratory study. Unpublished Ph.D. Thesis, University of Pittsburg.

RUSSELL, Harold E. & BENDIG, A.W. (1953) An investigation of the relation of student ratings of psychology instructors to their course achievement when academic aptitude is controled. Educational and Psychological Measurement, 13, 626-635.

SCHUH, Allen J., CRIVELLI, Michael A. (1973) Animadversion error in student evaluations of faculty teaching effectiveness. Journal of Applied Psychology, 58(2), 259-260.

STARRAK, J.A. (1934) Student rating of instruction. Journal of Higher Education, 5, 88-90.

STEWART, Clifford T. & MALPASS, Leslie, F. (1966) Estimates of achievement and ratings of instructors. The Journal of Educational Research, 56, 347-350.

SMELTZER, C.H. & HARTER, R.S. (1934) Comparison of anonymous and signed ratings of teachers. Educational Outlook, 8, 76-84.

SULLIVAN, Arthur M., SKANES, Graham R. (1973) Validity of student evaluation of teaching and the characteristics of successful instructors. Communication devant être présentée au congrès de l'APA à Montréal.

TOLOR, Alexander (1973) Evaluation of perceived teacher effectiveness. Journal of Educational Psychology, 64, 98-104.

TREFFINGER, Donald J. & FELDHUSEN, John F. (1970) Predicting students' rating of instruction. Proceedings, 78th annual convention, APA, 621-622.

VOEKS, Virginia W. & FRENCH, Grace M. (1960) Are student ratings of teachers affected by grades. Journal of Higher Education, 31, 330-334.

WALKER, Billy Doyle (1968) An investigation of selected variables relative to the manner in which a population of junior college students evaluate their teachers. Thèse de doctorat inédite, University of Houston.

WEAVER, Carl H. (1960) Instructor rating by college students. Journal of Educational Psychology, 51, 21-25.

WHERRY, R.J. (1952) The control of bias in ratings. VI. A comparison of various rating methods. PRB Report 921, Personal Research Bureau, Pers. & Res. & Proces. Div., Adjudant Gen. Off. Dept. of the Army.